数字人文视阈下明代科举进士群体时空网络结构分析*

邓君 孙绍丹 王阮 宋先智 宋雪雁
吉林大学管理学院 长春 130022

摘 要：[目的/意义]数字技术更新迭代，蓬勃发展，藉由数字智能化工具为人文研究提供新的研究视角和范式。[方法/过程]通过检索中国历代人物传记资料库(CBDB)中明代进士精英群体数据，借助社会网络分析(SNA)和地理信息系统(GIS)方法，绘制进士群体时空网络图。[结果/结论]从明代科举进士群体的时空演变规律、亲属关系网络及社会关系网络三个层次，"遥读"明代进士精英群体背后所折射的历史现象，以帮助人文学者从纷繁复杂的信息中高效地梳理史实资料，精准把握研究命题，更好地以鸟瞰视角探讨明代进士群体的整体时空分布规律及网络结构，彰显数字技术在数字人文领域中所隐藏的潜力。

关键词：数字人文 明代科举进士 时空网络结构
分类号：G270
DOI：10.13266/j.issn.0252-3116.2020.17.001

数字化浪潮下，人文活动在技术赋能下跨界融合，借助数字技术耕读人文知识，变革传统的人文研究范式，赋予人文研究新的活力。数字人文由人文计算演化而来，J. Unsworth 认为人文计算是由高效计算和人文交流需求决定的一种表示实践方式，一种知识建模形式、一种知识推理途径和一个本体约束[1]。数字人文拓宽了人文计算的外延和内涵，从多维度、广视角增加了研究的深度和视野，其意义在于通过智能化工具助力人文学者界定概念，解决人文问题[2]，回归人文本质。在大数据时代，数字化文献激增，数据资源类型多样，亟待技术助力挖掘数据背后所隐藏的人文内涵。数字人文研究离不开数据支撑，如文献数据库、领域专题数据库的构建。数据在人类认识历史、记忆和文化等方面发挥透镜作用，透过"冰冷"数据可以将人类历史文化"鲜活"地呈现在人们眼前[3]。

明代是中国科举制度空前完备的鼎盛时期，选官制度严格，科名等级森严。科举制度为朝廷输送了大量精英人才，形成"非进士不得入翰林""非翰林不得入内阁"之说。进士作为社会精英群体，与社会历史发展休戚相关。中国历代人物传记资料库(China Biographical Database Project, CBDB)[4]以宋元明人物传记资料索引及《清代人物生卒年表》、正史列传等史料、文集之祭文等文学作品以及官方文书为主要数据来源，将人物信息细化入库，建置关系型数据库。截止到2019年4月，共收录约427 000人的传记资料，其中囊括明代科举－进士(笼统)、科举－武举进士和科举－正奏名进士共计28 311位。本文以此为数据源，依托社会网络分析(SNA)和地理信息系统(GIS)方法，立足数字人文视阈，遵循数字人文研究范式，探讨明代科举进士群体网络关系及时空演变规律，运用技术工具将该群体时空分布架构可视化呈现出来，"遥读"明代进士群体时空网络关系，为人文学者细粒度研读明代进士史料提供新的视角，重拾人文研究"温度"，激发人文研究焕发新活力。

1 相关研究

数字人文是一个方兴未艾的学术新领域。公认源于 R. Busa 在1949年用计算机为阿奎那作品编制拉丁文索引[5]，这是首次将信息技术应用于人文领域。21世纪初，数字人文组织联盟建立、刊物出版，将数字人文研究推向高潮。如今，在大数据和人工智能时代，数字技术更新迭代，蓬勃发展，为人文研究提供了更大

* 本文系国家社会科学基金项目"数字人文视角下历史档案资源知识聚合与知识发现研究"(项目编号：19BTQ102)研究成果之一。
作者简介：邓君(ORCID：0000-0002-3291-7193)，教授，博士生导师；孙绍丹(ORCID：0000-0003-2927-0971)，博士研究生，通讯作者，E-mail：1651873190@qq.com；王阮(ORCID：0000-0002-5445-4943)，博士研究生；宋先智(ORCID：0000-0002-7499-2431)，博士研究生；宋雪雁(ORCID：0000-0003-1955-6685)，教授，博士生导师。
收稿日期：2020-03-23 修回日期：2020-05-30 本文起止页码：4-17 本文责任编辑：易飞

的契机,更能从宏观视角"远读"人文研究中的现象、解构内在机理。自 F. Moretti 提出"distant reading"概念以来[6],量化研究就在人文领域生根发芽,如 M. Schich 等采集 15 万多位著名人物出生地与去世地信息,利用网络工具和复杂理论量化文化史研究[7],是历时性研究的典范;数字人文学者 M. L. Jockers 在其著作"Macroanalysis"中首次剖析文本挖掘技术在文学领域的应用[8],从宏观视角解读人文研究中的现象。另外,学者们在大量语料库基础上利用计算技术论证作品归属[9]、作者写作风格、作者情感倾向[10-11]等问题,也有通过机器学习算法挖掘古籍中实体信息[12]及发现古籍中隐藏的新现象[13]。可以看出,计算技术助力人文研究以"鸟瞰"方式探勘人文语料,进而发现新问题,开拓人文研究视野,变革传统的仅以思辨式方式研究人文问题的范式。这样的"远读"策略可以为人文研究提供新思路,更直观地再现人文研究全貌。

纵观目前的数字工具及方法, SNA 和 GIS 方法为数字人文研究提供了可能性。SNA 在人文领域应用维度多样,在文学作品中人物关系构建方面,人文学者 F. Moretti 利用《哈姆雷特》中人物关系,诠释莎士比亚悲剧主题"维护君主正统性"[14];赵薇用 SNA 方法梳理《大波》三部曲中人物关系网络[15];魏会洋等用 SNA 分析《白鹿原》中的人物关系[16],可以看出,SNA 方法适用于梳理文学作品中复杂的人物关系。在历史研究中,C. Wetherell 首次将 SNA 方法引入史学领域,探讨其如何促进对历史血缘关系的理解[17];也有学者分析历史人物群体的整体和个体网络特征,试图厘清历史人物的网络关系,如严承希等探讨宋代政治网络中的整体网络分布特征、核心人物地位与结构拓扑以及网络时序关系演化模式[18];J. NOVAK 等开发了 histoGraph 工具,旨在助力历史学家对人、地点和事件之间的关系进行发现和历史分析[19]。在艺术人文领域,L. Weixuan 构建了历史绘画中艺术思想网络[20],进而挖掘其深层内涵。GIS 虽然属于地理学方法,但在史学和文学研究中也有所应用。20 世纪 90 年代,A. K. Knowles 和 I. N. Gregory 将 GIS 引入到史学研究中[21-22],形成 HGIS 概念,勾勒了史学研究愿景图,变革了史学研究范式,为史学研究提供空间化视角。国内最典型的历史 GIS 当属 CHGIS,该项目试图建立公元前 221 年-1911 年间连续的中国历史行政单位数据库,为研究者提供数据平台和空间分析模型。此外一些学者针对具体历史事件构建一系列历史 GIS,如丝绸之路历史 GIS 平台[23]、华夏家谱 GIS 平台[24]及中国古籍 GIS[25]。在研究文献方面,有学者论证了 GIS 与中国历史研究的关系[26],并从地理与历史双重视角出发,设计出一种基础的、通用的历史 GIS 数据模型[27],也有学者对诗人地理分布进行可视化呈现与阐释[28]。

综上所述,除了计算算法在人文领域的应用之外,SNA 和 GIS 方法也受到研究者青睐,但是就本文的研究问题"明代进士群体",尚未看到这两种方法的应用。虽然史学界对明代进士群体的研究不胜枚举,但主要以定性化方式为主,缺乏量化分析。研究范式以阐释、总结、归纳为主,且以小范围、小视角方式切入进士群体研究,如以某一省、县研究其空间分布特征及影响因素[29],以某一进士家族为例分析其兴衰史及背后之因[30],缺乏对明代整个进士群体的全貌分析;在进士数量时空分布上,现有的研究中往往以表格描绘各时期进士数量,以文字论述进士分布地点[31],缺乏视觉上的直观性,不利于发现问题;对于进士网络关系的研究,往往是文字形式阐述进士之间的亲属关系[32],对社会关系的探讨少之又少。因此,本文将 SNA 和 GIS 引入到明代科举进士群体研究中,从时间、空间和网络视角探讨该群体背后所折射的历史现象,为人文学者深入解读现象背后之因提供一种分析路径。

2 研究方法

2.1 社会网络分析(SNA)

社会网络被界定为一种或多种关系连接而成的点集[33],通过中心性、平均度、聚类系数、平均路径长度和网络密度等衡量网络结构特征。SNA 主要从网络视角挖掘人文领域人物血缘、亲属、社会关系等,并研究参与者在社会场景中交织而成的社会关系型构[34]。本文的进士属于明代精英群体,用 SNA 方法研究该群体的网络拓扑结构及网络特征,能够从网络视角明晰群体的社会关系属性,梳理人物之间的多重网络关系,从关系中发现潜在现象及特征,为人文学者深入解读背后的内在机理提供参考。

2.2 地理信息系统(GIS)

GIS 隶属于地理学研究范畴,在数字人文热潮下,也是备受关注的方法。GIS 将复杂的空间技艺融入到人文研究中,更丰富地诠释了人文内涵,拓展了研究视角。明代进士群体出生地差异显著,空间分布是随时间动态演化的,因此引入 GIS 方法可以从空间视角直观呈现进士的区域分布态势,以地图叙事方式发掘距离因素对进士分布的影响,为人文学者解读现象提供思考路径。

3 明代科举进士群体时空分布规律探析

3.1 数据来源及预处理

本文从 CBDB 数据库检索明代进士信息，筛选科举－进士(笼统)、科举－武举进士和科举－正奏名进士，共计 28 311 位进士信息，预处理后得到 17 408 位进士(见表 1)。导入中华文明之时空基础架构－WMTS 服务框架[35]，抽取明代疆域地图作为底图。

表 1 明代进士群体区域分布数量

区域	数量(位)	区域	数量(位)
山东	1 127	湖广	1 016
京师	1 304	广西	146
山西	782	云南	70
河南	1 034	贵州	14
南京	2 996	四川	912
浙江	2 805	陕西	651
江西	2 280	广东	614
福建	1 657		

3.2 明代科举进士群体时间序列分析

明朝共有 17 位皇帝，从图 1 可以看出永乐－洪熙和万历－泰昌年间被归一。永乐二十二年(1424 年)八月，明仁宗登基，次年(1425 年)改年号为洪熙，因洪熙帝在位仅 10 个月，将其和永乐帝归在一起。泰昌帝在位仅一个月且执政期间并未有进士，因此将其归入到万历年。图 1 显示了明朝 4 个进士峰值点(永乐－洪熙年、成化、嘉靖和万历－泰昌)和 2 个低谷点(建文和天启年)，宣德至天顺年间进士数量较为稳定。获取大量进士数据，借助图形直观地呈现明代进士录取额的高峰和低谷期，针对性地梳理史实资料，探讨明代科举制度分期发展阶段性变化特征，多维度考量分期进士录取情况，耕读史料并深究其因。

3.3 定量视角研讨明代“南北分卷”制度影响

明洪武三十年(1397 年)“南北榜”事件发生，南方进士占绝对优势而引发北方人不满，自此分卷制度逐渐建立，在此定量化明代进士数据，以揭示分卷制度对南北进士录取额的影响。图 2、3、4 分别是南卷、北卷、中卷各省进士数量时间序列图。总体上看，南北卷制度实行对南方进士数额并未有太大影响，南方进士仍远高于北方，约是北方进士数额的 2 倍，但由图 3、4 可知，分卷制度实施初期(永乐－洪熙年)，南卷进士数量约是北卷进士数量的 7 倍，说明实行初期南卷依旧占大比例，而从宣德年始，南卷进士数量是北卷 2 倍左右，并一直保持此比例，这其中缘由与人口流动、科举政策等因素相关。究其实质，分卷制度着实给北方人提供了爬升社会阶梯的机遇，其进士数量在波动增加，幅度与南方进士数量演变趋于一致。中卷中四川省从永乐年伊始，持续在中卷中保持领先优势，但最高峰嘉靖年也只有 250 位进士，其余省因人口稀少、经济发展迟缓，进士数量在 50 以下。

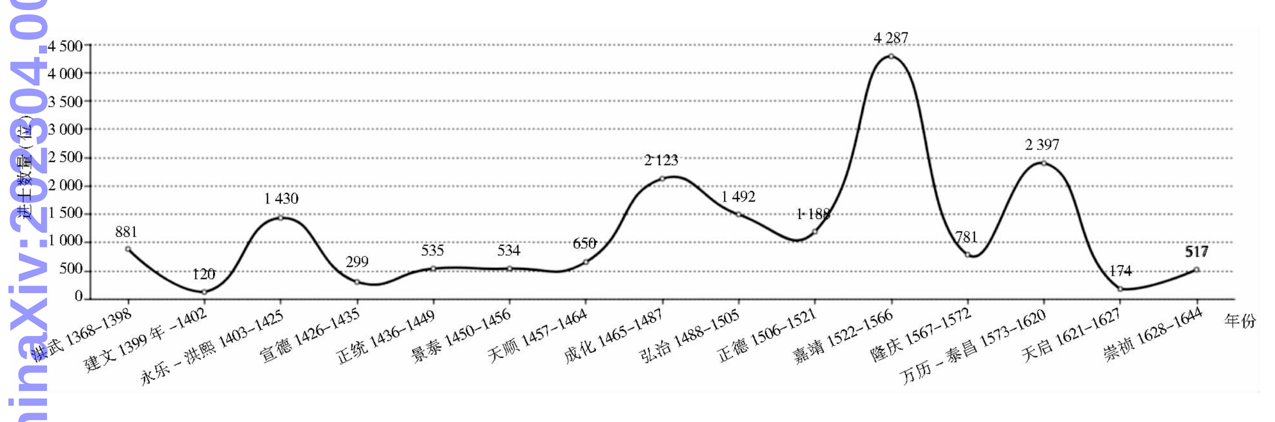


图 1 明代科举进士数量时间序列

3.4 明代各省进士数量演变主题河流图

从图 5 可清晰辨别某一时期某省进士数量多寡及整个明代各省进士演变历程。图 5 明代各省进士数量演变趋势与图 1 整体进士数量变化几乎吻合，且显示出 4 个高峰点与 2 个低谷点。其中，永乐年江西进士数量占比最多。嘉靖年是各省进士数量最多的时期，史学界对进士分期差异化影响因素研究并不多见，详尽原因还需考证史料深入研读。总体来看，南京、浙江、江西、福建河流覆盖面最宽，进士数量名列前茅，与图 6 进士分省总体数量柱状图遥相呼应。主题河流图可以直观地凸显问题表象，从现象到本质，从表层到深层助力人文学者深化研究，节省时间和精力，更有效率地挖掘问题背后的深层次历史内涵。

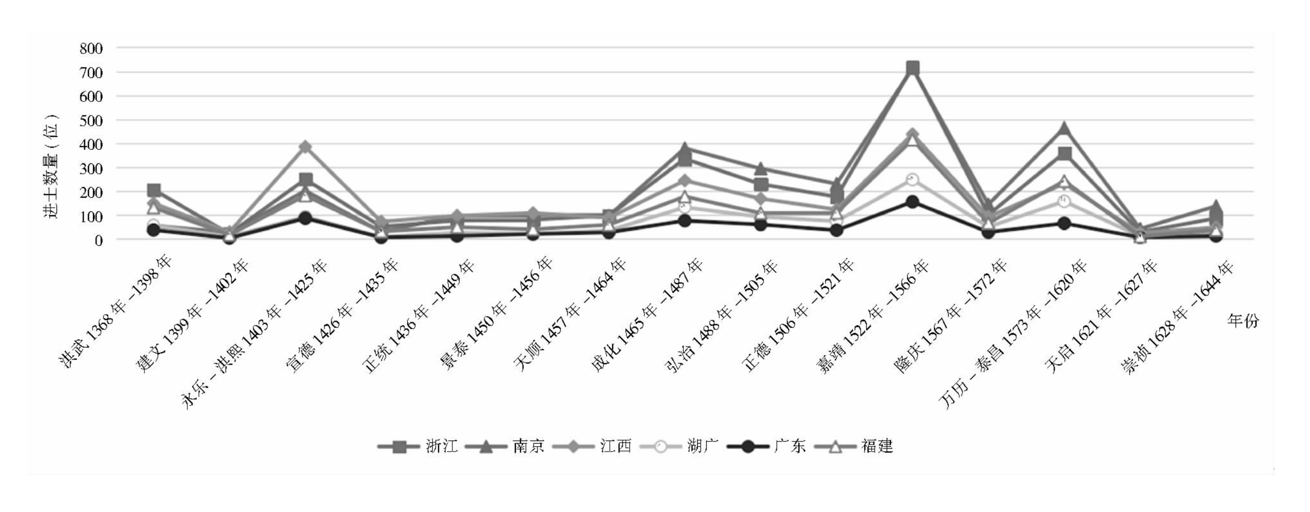


图 2 “南卷”中各省进士数量时间序列

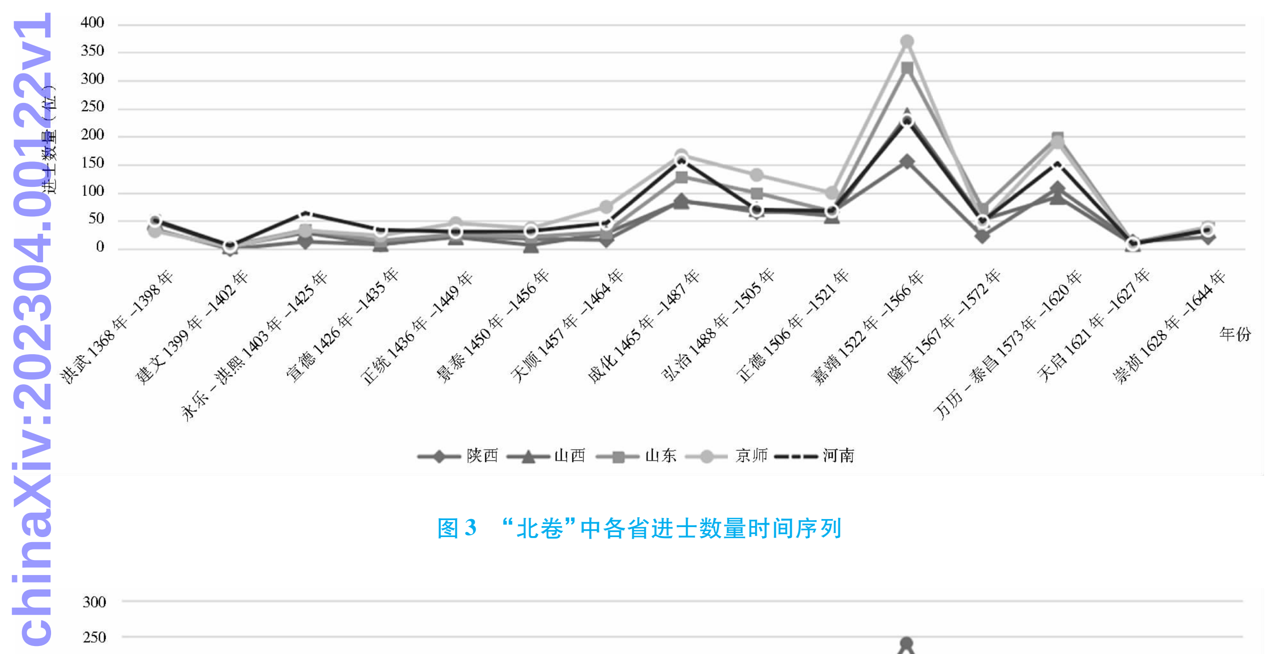


图 3 “北卷”中各省进士数量时间序列

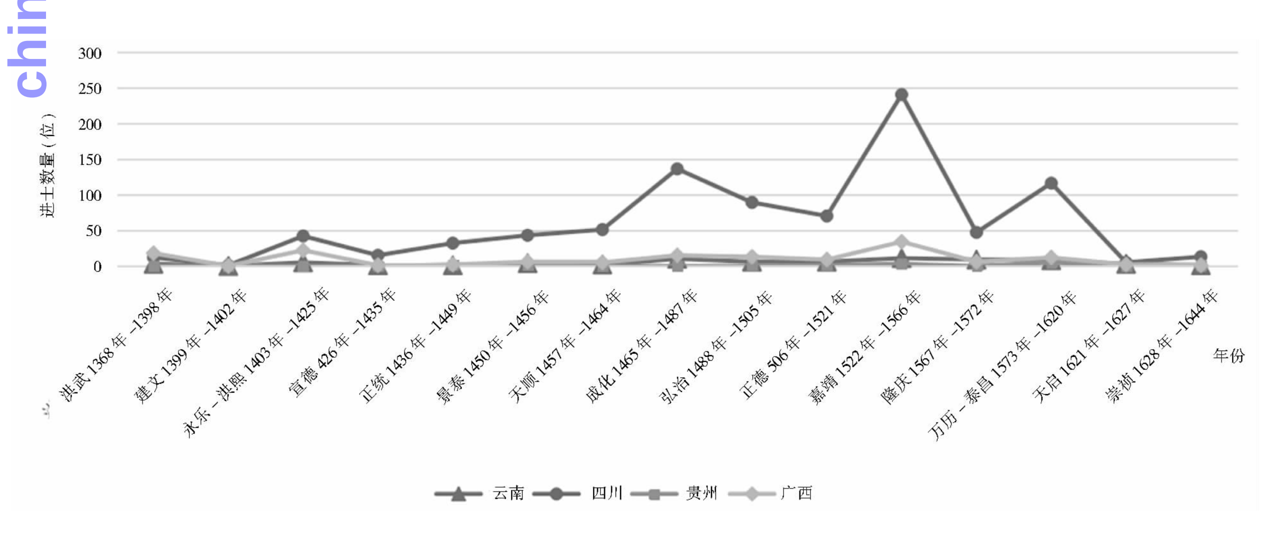


图 4 “中卷”各省进士数量时间序列

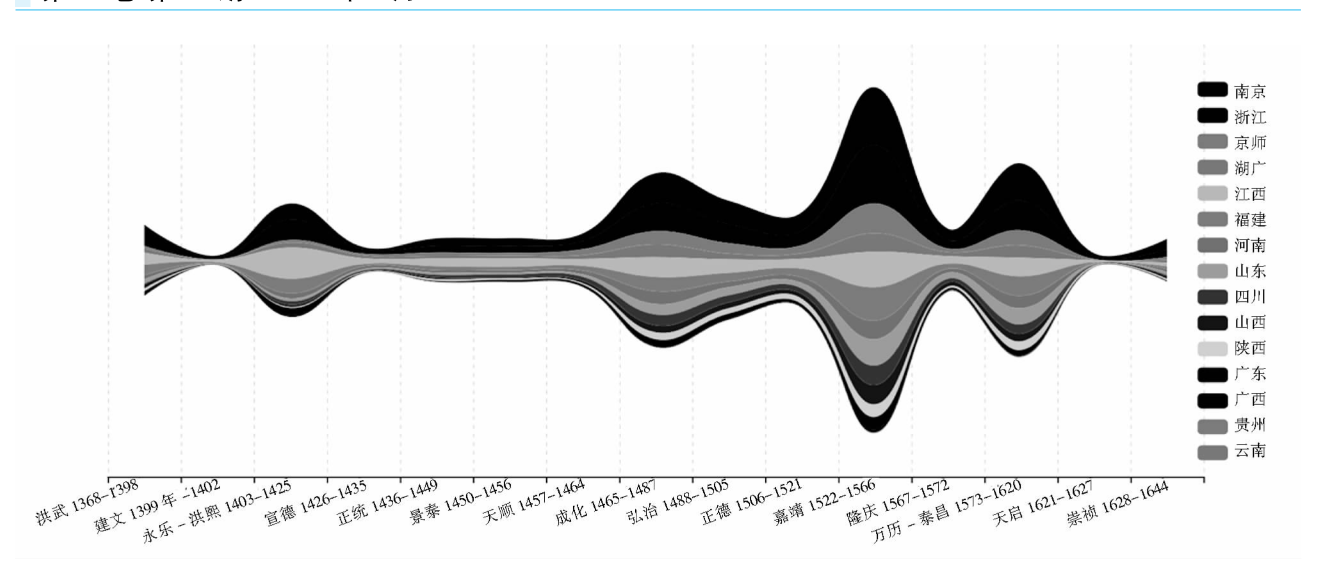


图 5 明代各省进士数量演变主题河流图

3.5 明代各省、府、县进士空间分布

空间是历史事件演化进程中不可或缺的因素，历史事件在空间地图上的映射承载了历史记忆，从空间视角寻绎潜在的历史现象方可催生新的研究范式。

据图 6 可将进士群体区域分布数量划为三梯度，即南京、浙江、江西、福建省位于第一梯度，山东、京师、山西、河南、湖广、四川、山西、广东位于第二梯度，广西、贵州、云南位于第三梯度，进士数量依次递减。进士区域分布差异化研究成果丰富，几乎每个省份都有史学学者详尽阐释过[36－37]。

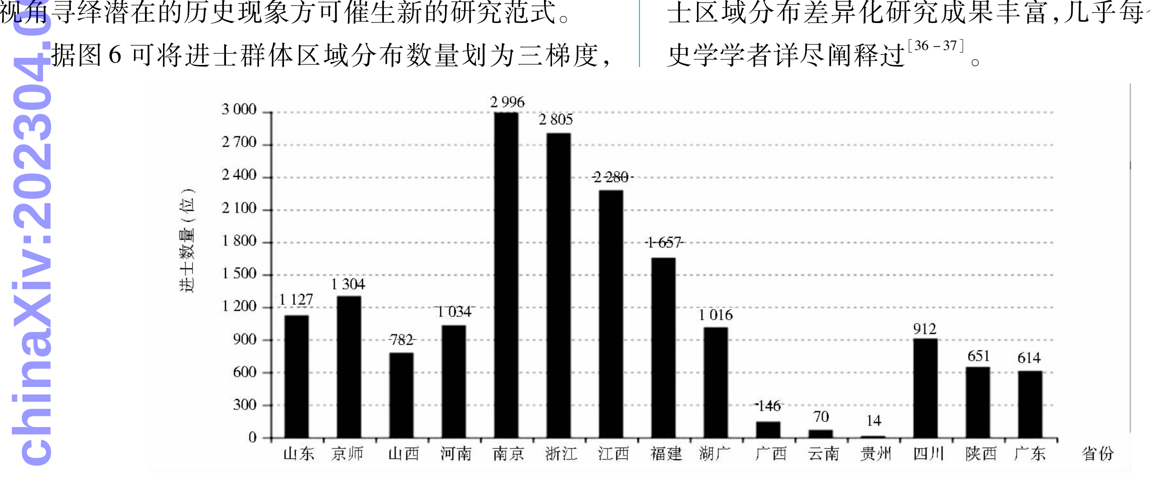


图 6 明代各省进士数量空间分布

用 GIS 技术将进士分布进行空间可视化，更细粒度呈现明代各省、府、州、县进士分布情况。由图 7 可以看出，明代各省进士分布极不均衡，东南沿海区域较多，东北以北直隶为核心向四周扩散分布，西南和西北区域极少，由此方可论证进士荟萃之地的必然和偶然因素及空间分布差异化的影响因素，这类因素可通过获取历史数据，借助 GIS 图层叠加功能，在地图上重现要素对进士群体分布的影响，如人口分布、进士配额等映射到地图上，从中探索隐匿规律，为史学研究提供新的生长点。

以南京、浙江、江西、福建为例，更细粒度绘制四省府内进士分布情况。由图 8、9、10、11 可以清晰直观地看出，进士数量分布较多的府包括浙江东北部沿海嘉兴府、湖州府、杭州府、绍兴府、宁波府、台州府；南直隶东南部苏州府、松江府、常州府、镇江府、应天府、徽州府；江西中北部吉安府、临江府、抚州府、南昌府；福建东部沿海福州府、兴化府、泉州府、漳州府，与何炳棣先生从史料中归纳的科甲鼎盛区域如出一辙[38]。从浅层空间分布图来看，进士多的区域主要集中于东部沿海各省、府，可以推知经济因素对进士数量影响深远。从深层内涵剖析，这 4 个省内各府进士数量分布极不均衡，南直隶南部与浙北接壤区域进士数量较多，且各省内进士数量居多的府之间距离较近，说明空间距离因素对进士录取率会有明显影响。从图 10 得知江西作为内陆地区，进士数量却位居前列，具体之因有待进一步考察。通过图形可以直观再现明代各省府进士空间分布态势，突破传统的人文研究范式，为人文研究带来便捷。

注:图例中圆形大小表示地图中柱状图高低,圆形越大表示地图中柱状图越高,进士数量越多,反之亦然

图7　明代进士群体空间分布

针对空间因素可以反思明代道路交通因素是否会对进士流动产生作用。齐普夫“省力法则”在进士分布态势图中也可以清晰地被映射出来。“就近原则”可能会影响文化传播与交流,进而变革固定领域范围内人员的知识结构,在文化熏陶下,甚至可能会推动平民阶层向上流动,改变其命运,这就是数字人文的潜力所在。数字技术能为人文研究提供新的研究命题和研究范式,探索未知因素,多角度深化史学研究,从空间视角审视人文研究,以“问题”为导向,推动人文研究向纵深方向发展。

为进一步从县级视角窥探进士数量分布情况,列出明代进士数量在100位以上的县(见表2)。由表2可以看出,明代进士主要集中于福建、浙江、江西和南直隶省,其中河南省祥符县有102位进士,广东南海有103位,此结果可以驱动史学界学者多角度论证县级进士萃集之因。从图12可以看出,进士数量较多的县总体上呈现“聚类”分布特征,即空间上集聚效应明显,各省内进士汇集之地空间距离较近,据此可以对空间距离做量化分析,精要严谨地加以斟酌。

图8　南直隶各府进士数量空间分布



图9　浙江各府进士数量空间分布

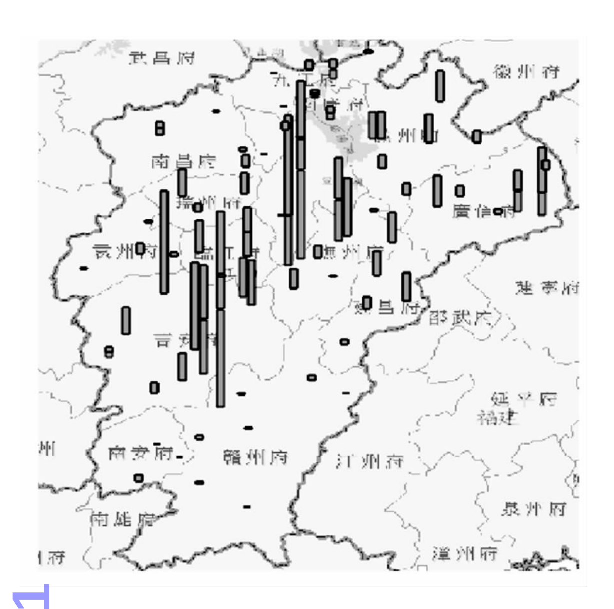


图 10　江西省各府进士数量空间分布

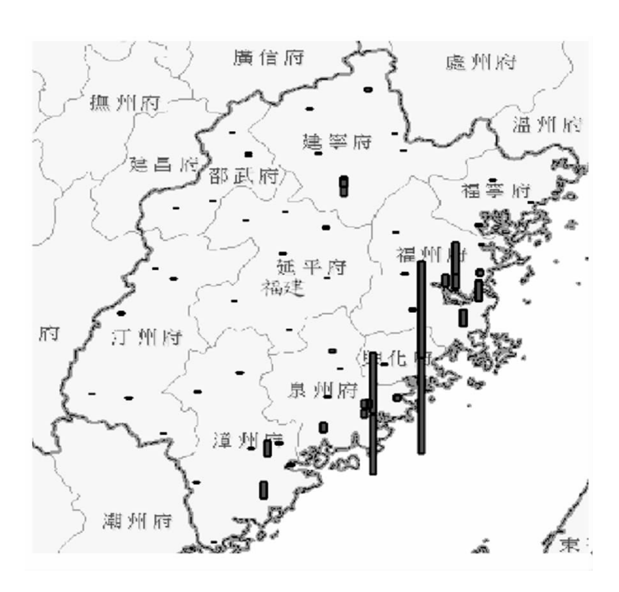


图 11　福建省各府进士数量空间分布

表 2　明代进士数量最多的 24 个县(进士数量 > 100)

县	数量(位)	县	数量(位)	县	数量(位)
福建莆田	381	江西吉水	161	南直隶常熟	143
浙江余姚	306	南直隶昆山	157	浙江仁和	141
福建晋江	239	南直隶长州	153	南直隶武进	139
南直隶华亭	198	浙江慈溪	152	南直隶无锡	136
浙江鄞县	191	南直隶吴县	149	浙江钱塘	110
福建闽县	188	江西南昌	146	广东南海	103
江西安福	169	浙江山阴	146	河南祥符	102
江西丰城	165	江西泰和	145	南直隶歙县	101



图 12　明代进士数量最多的 24 个县分布

4　明代科举进士群体网络关系结构研究

4.1　亲属关系

通常在古代社会,人是基于血缘、地缘和宗教共同体并以群落形式聚集生存。

弗里德曼认为,在研究中国文化和社会时一定要考虑和中国人生活密切相关的谱系和亲缘问题[39]。明代进士精英群体家族亲属关系网值得学者深度挖掘探索。郭培贵认为五代直系亲属内有两名以上的可称之为进士家族,并研讨明代进士家族对政治和社会的影响[40]。人文学者从定性视角对进士家族概况、兴盛之因、政治影响等进行了相关研究[41]。本文从定量视角探求直系祖先对进士入仕率的影响和中国第一科举家族——常州庄氏家族网络拓扑,更直观地彰显进士家族群体拓扑结构,为人文研究者提供方法参考。

4.1.1　明代进士群体直系祖先有官职对入仕影响

由图 13 可知,直系祖先中无官职的比例占 45.74%,有官职的比例 54.28%,总体来看,各占据一半左右,说明进士群体在向上流动过程中,非官僚阶层占据了近乎一半比例,给予了寒微人士攀爬社会学术阶梯的机遇,对明代社会流动和社会结构变迁起了一定的作用。直系祖先具有官职某种程度上会由于社会关系、家庭文化熏陶等因素对进士提升产生影响。

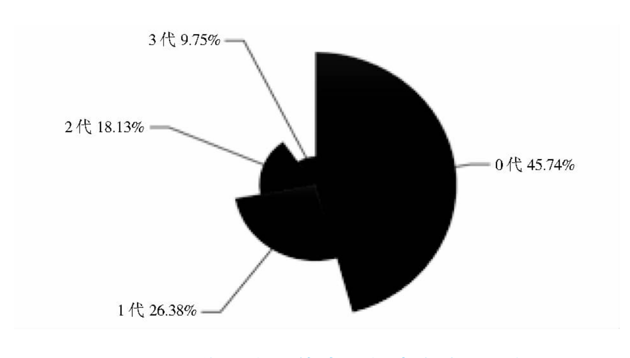


图 13　明代进士群体直系祖先有官职比例

从图14可以看出明初直系祖先有官职的比重并不大，从永乐年伊始，直系祖先官职比例骤升，到正德、嘉靖年官职比重最高，约达到65%，对比明代进士数量时间序列图1，可知嘉靖年间进士数量最高，达到4 287，据此推测进士数量之多和直系祖先有官职可能存在某种联系，进一步验证了除政策文化地理环境人口因素外，直系祖先有官职对进士入仕比例有所影响。

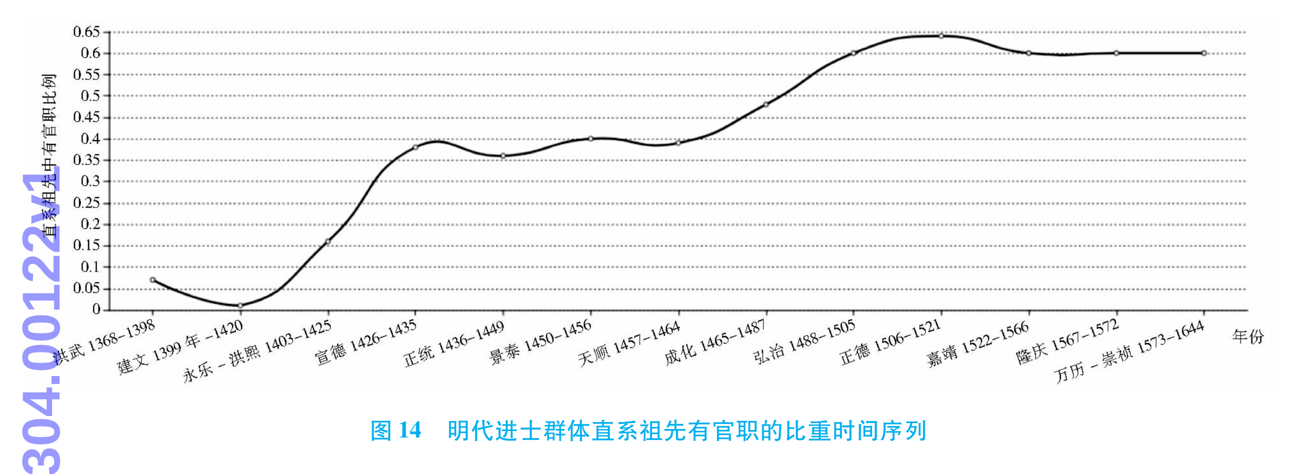


图 14　明代进士群体直系祖先有官职的比重时间序列

4.1.2　中国科举第一家族“常州庄氏”世系关系呈现

家族树是一种最直观的呈现方式，在数据量庞大时绘制家族树，易于梳理家族人物关系，直观呈现血脉亲属关系，提高研究效率，并可结合家谱、传记等资料检视家族枯荣浮沉之史。人文领域对此应用甚少，往往定性枚举。这里仅以CBDB收录的明代庄氏家族数据为例，以树状图形式呈现(见图15)。明清时期，中国科举第一家族“常州庄氏”科甲鼎盛，共中举79人，中进士35人，其中第八代同祖兄弟庄起元和庄廷臣双双考中进士，仅以此为切入点，在CBDB数据库中查询，简要绘制树状关系图，借此参考。亦可结合庄氏族谱，扩展数据来源，搜集家族中部分人物关系，以图谱形式展现。同时综合姻亲关系，查阅庄氏家谱，全视角凸显家族世系关系。借助数字工具表征家族世系关系较传统方式更加便捷、直观，其关系可一目了然。

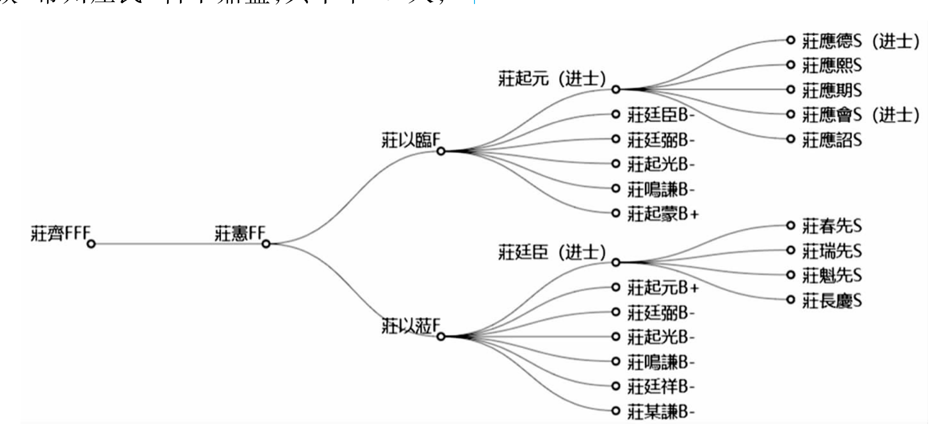


图 15　“常州庄氏”世系关系树状图(部分)

4.2　社会关系——以明末“东林进士八君”社会关系网络为例

《中国历史大辞典》“东林书院”条释义：东林八士大夫议论朝政，品评人物，参与政治、反对空谈，朝士亦遥相应和，号称东林八君[42]，此“东林八君”即顾宪成、顾允成、高攀龙、安希范、刘元珍、钱一本、薛敷教、叶茂

才，八位人物均为明代进士。这里仅以“东林八君”小进士群体为核心，解构其社会关系网络拓扑。CBDB 中社会关系类型包括朋友、学术、政治、著述关系等。从 CBDB 中按 ID 号检索“东林八君”，距离设置为 2，共检索到人物 635 位、关系 2 170 个，构建“东林八君”社会网络关系知识图谱。

为更清楚地从整体层面理解明代“东林八君”社会关系网络结构特征，引入网络分析指标对其进行整体性网络结构分析，通过计算得出网络平均度为 5.153，平均聚类系数为 0.277，平均路径长度为3.392，密度为 0.008，网络直径为 5。平均聚类系数较大，平均路径长度较小，即平均通过 3－4 个人就可互相连通，具有小世界效应的显著特性[43]。

在双对数坐标轴中，以散点图形式绘制双对数坐标图 16，得到拟合直线 $y = 0.320\,3x - 1.409$，$R^2 = 0.852\,1$，R^2 越接近于 1，越呈现出明显的线性趋势。由图 16 度分布图看出，该网络符合幂率分布特性，是一个无标度网络[44]，即各节点之间连接状况（度数）具有严重的不均匀分布性，网络中 Hub 节点拥有极多连接，而大多数节点只有少量连接，少数 Hub 点对无标度网络运行起主导作用，且显现出长尾特征。明代“东林八君”社会关系网络与当代社会媒体网络结构存在相似之处，侧面彰显了数字人文视域下 SNA 方法的潜力所在。

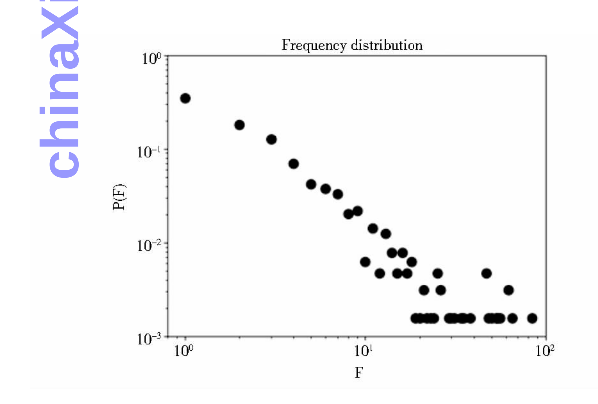


图 16　度分布图（双对数图形转化）

4.2.1　“东林八君”社会关系网络中核心人物分析

度中心度指的是与该节点邻接节点的个数，是衡量网络节点核心地位的指标之一。在人类交往方面，较高度数的点在某种程度上是“核心”，在交往活动中具有很强的控制力。图 17 表示明末“东林八君”社会关系网络，按中心度由大到小顺序对人物进行排序，用节点字体大小区分人物。表 3 是根据网络结构中节点度大小排序的前 30 位人物。由图 17 和表 3 可知，中心度在 30 以上的人物有钱谦益、姚希孟、罗钦顺、李维桢、王世贞、邹元标、徐学谟、顾起元、王守仁、陆树声、顾宪成、陈懿典、徐阶、毛宪、叶向高、薛瑄、冯丛吾。首先，这些都是进士及第身份之精英群体，在当时名声大噪。以钱谦益、邹元标、顾宪成、叶向高、高攀龙等为代表的东林学派，以儒家经史著述为讲学内容，推动了史学思潮的发展。其次内阁辅臣身份之人也名列其中，如姚希孟、陆树声、徐阶、叶向高，针砭时弊，建言献策，稳定政局。另外也不乏文坛领袖及思想家，如以薛瑄、冯丛吾为代表的“关中学派”、王守仁“阳明心学”、罗钦顺“气学”等，推动明朝哲学思想传播演进。这些核心人物某种程度上诠释了“东林八君”在当时社会的影响力之大，囊括了诸多名流之辈。

4.2.2　“东林八君”社会关系网络小群体现象分析

采用 M. E. J. Newman 等提出的测量网络社群聚类程度的模块度算法[45]，用以衡量聚类效果，模块度值域范围[－1，1]，越接近于 1，代表网络社群结构强度越大。一般认为，模块度值 >0.3，代表网络具有明显的社群结构，该网络模块度值为 0.541，说明此网络具有良好的社群结构，且结构强度较强。由图 18 和表 4 可以看出，划分为以罗钦顺、姚希孟、李维桢、钱谦益、王世貞、陆树声、徐学谟、邹元标、薛瑄为核心的 9 类社群，节点大小代表度值，连线表示相关关系。根据表 5 列出的 9 类社群人数可以看出，除第一类人物领先之外，其余类人数差别并不显著。

群体中影响力最大的是以罗钦顺、王守仁和毛宪为代表的第一类社群，该社群内节点度值在 10 以上的有 19 个，总人数最多。该类群体以思想家、哲学家为主体，以王守仁、湛若水为首，素有“王湛之学”之称，二者弟子遍布各地。罗钦顺发展了程朱理学，创建“气学”。严嵩虽然被列为奸臣，但位高权重之前广结王阳明等名流，学识渊博，剖经析义，在文学层面造诣颇深。该类社群人物影响力最大，某种程度上揭示了程朱理学、阳明心学在明代“东林学派”社交圈内被发扬传承范围之广。

以姚希孟、顾宪成、高攀龙、赵南星为首的第二类社群，人数居第二位，核心人物度值相对较高。作为东林学派领袖，追随者众多，说明东林学派在明代声望之高。此类中还有诸多东林知名人士，如汤显祖、杨涟、安希范、叶茂才、李三才、顾允成、于孔兼、钱一本、薛敷教、杨廷筠、刘元珍等，且“东林八君”聚类于此，侧面验证了模块度算法的准确性。由图 18 也可以发现，第二类部分人物交融在以邹元标为核心的社群中，因此第八类可以和第二类归为一类。

图 17 明末“东林八君”社会关系网络(距离为 2)

表 3 “东林八君”中心度排序(前 30)

Label	degree	Label	degree	Label	degree
钱谦益	84	顾宪成	47	罗洪先	25
姚希孟	65	陈懿典	47	張鼐	25
罗钦顺	62	徐阶	38	王锡爵	25
李维桢	62	毛宪	35	高攀龙	24
王世貞	56	叶向高	34	严嵩	23
邹元标	54	薛瑄	31	沈鲤	22
徐学谟	53	冯从吾	30	焦竑	21
顾起元	50	吕枏	29	汪道昆	21
王守仁	48	姜宝	26	茅坤	20
陆树声	47	湛若水	26	于慎行	19

以陆树声、徐阶、姜宝为首的第六类社群，节点度值在 10 以上的有 17 个，这类社群人数虽然落后于前五类，但所涵盖的人物影响力却不逊色于第一类。除此之外，此类群体还囊括了赵志皋、张居正等内阁首辅，著名谏臣杨继盛。此群体主要以政治人物为代表，利益关系错综复杂。

第三类和第五类社群呈现出文学家与政治家融合

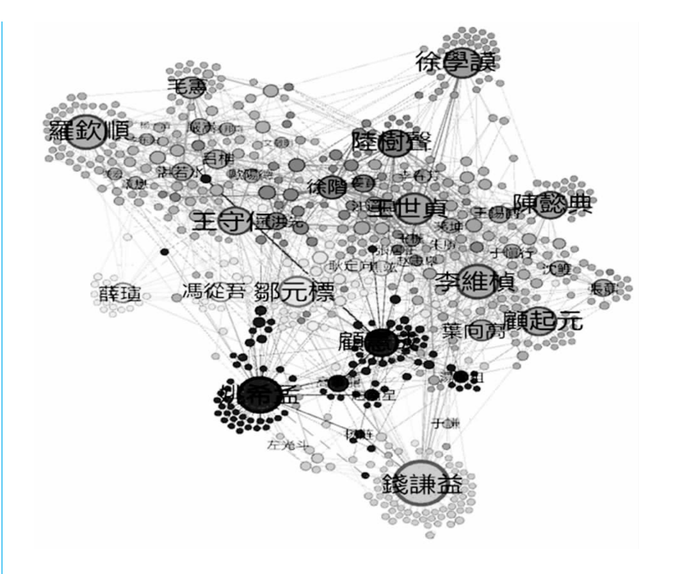


图 18 “东林八君”社会关系网络模块

特征。晚明文坛盟主李维桢、金石家书法家顾起元、以王世贞、李攀龙为代表的明末文学流派“后七子”、画家陈继儒，也有内阁首辅叶向高、李春芳、王锡爵、沈一贯、王家屏等。

以钱谦益为首的第四类社群和以徐学谟为首的第七类,呈现相似分布特征,核心人物度值最高,但与之关联的人物影响力却较低。

以薛瑄、冯从吾为首的“关中学派”被聚合为一类,冯从吾是将程朱理学和陆王心学融合的集大成者,也是东林学派在西北的领袖,因此和邹元标距离较近。

此外,在小群体中,存在多维社会关系交织融合现象。历史人物并非孤立存在的个体,其关系纽带是传递信息知识的显著渠道。既可从微观层面解构个体人物之间互动关系,也可从宏观视角剖析社会网络结构本质。“东林八君”社会关系网中包罗的人物关系具有多维属性,多种社会关系交织融合,如著述关系、政治关系、朋友关系等,暗示着东林人物之间具有复杂多变的社会关系,对传播东林学派思想文化起着不容忽视的作用。

表 4 “东林八君”社会关系网络模块化代表人物度值

第一类	degree	第二类	degree	第三类	degree	第四类	degree	第九类
罗钦顺	62	姚希孟	65	李维桢	62	钱谦益	84	薛瑄
王守仁	48	顾宪成	47	顾起元	50	倪元璐	9	冯从吾
毛宪	35	高攀龙	24	叶向高	34	陳仁錫	7	曹于汴
吕柟	29	汤显祖	14	張鼐	25	于謙	7	楊爵
湛若水	26	赵南星	13	沈鲤	22	張溥	7	南軒
严嵩	23	罗大紘	10	于慎行	19	黄道周	5	南大吉
欧阳德	18	徐用检	8	李春芳	18	李流芳	4	李賢
张邦奇	16	邵宝	7	冯琦	13	沈守正	4	陳祚
章懋	16	孙鑛	7	许国	13	祁承爜	4	孫丕揚
第五类	degree	第六类	degree	第七类	degree	第八类	degree	degree
王世貞	56	陆树声	47	徐学谟	53	邹元标	54	31
陈懿典	47	徐阶	38	顾绍芳	7	焦竑	21	30
王锡爵	25	姜宝	26	陆粲	6	耿定向	18	9
汪道昆	21	罗洪先	25	顾章志	5	朱賡	16	6
茅坤	20	邹守益	18	祝允明	3	羅汝芳	15	6
冯梦桢	17	胡松	17	甘雨	3	張元忭	10	4
陈文烛	17	王畿	16	秦嘉楫	3	周汝登	9	3
赵用贤	12	聂豹	15	华叔阳	3	楊起元	9	3
王世懋	11	王樵	14	黄体仁	3	劉日升	9	3
潘恩	11	万士和	13				3	

表 5 “东林八君”社会关系网络模块类人数

群体	人数(个)	群体	人数(个)	群体	人数(个)
第一类	139	第四类	77	第七类	38
第二类	96	第五类	65	第八类	37
第三类	88	第六类	60	第九类	35

4.2.3 “东林八君”社会关系人空间分布态势

绘制“东林八君”社会关系网络人物空间分布图(见图 19),可以看出人物主要分布于三个区域,即南直隶南部、浙北及江西中部,以北直隶(京师)为核心周边区域和陕西关中地区。空间分布特征映射了东林学派思想传播的空间维度。东林学院坐落于南直隶无锡,顾宪成、高攀龙等诸多领袖皆是无锡人,以此为轴心扩散于周边区域,对照进士空间分布图 7,可以发现进士荟萃之地亦是“东林八君”社会关系人聚集之地,侧面验证了东林人社会影响力很大。

总体来看,“东林八君”社会网络符合小世界效应及无标度分布特征,接近于现代社交媒体网络结构属性特征。从“东林八君”社会关系核心人物、小群体现象及人物空间分布态势三维度阐明了该社会网络呈现的特征及所隐含的历史现象。从网罗的社会关系人来看,延揽诸多遐迩闻名之士,如多位内阁首辅及知名思想家等,他们在网络中扮演的角色意义显著,对整个东林学派甚至东林思想发展趋向起着不可磨灭的作用,凸显了社会网络分析的潜在价值。



注:图例中圆形大小表示地图中柱状图高低,圆形越大表示地图中柱状图越高,进士数量越多,反之亦然

图19　明末“东林八君”社会关系人空间分布

5　结论与展望

数字人文在大数据时代下生存空间得以拓展延伸,通过智能化工具可以打破技术与人文之间的壁垒。传统的人文研究范式,往往以手动整理人文数据、梳理、阐释、归纳、总结相关研究问题、细读史实资料,并结合独立思考探讨既有的史学问题,效率相对较低。而数字技术的运用,方可以鸟瞰方式提出新的研究问题,驱使人文研究从表层文本现象深入到背后潜在的本质内涵研究中,以“问题”为导向,以“人文”为本位[46],开拓创新,构思新的研究命题,更有效率地从海量史学资料中发掘内在机理,拓展研究思路。本文以明代进士精英群体为研究对象,引入SNA和GIS方法,在数字人文视阈下,从进士群体时空分布规律、亲属关系及社会网络关系三个维度发掘进士群体背后所折射的史学现象及本质:

(1)在进士群体时空分布规律中,以可视化图形表征明代进士时空分布态势;从明代进士的时间序列分布折线图可知,明代进士数量存在4个高峰点和2个低谷点,且在嘉靖年数量最多,这可以帮助史学者快速梳理史学数据,为人文学者多维度考量进士时序变化之因提供新思路。此外又从定量化视角探讨“南北分卷”制度下南北进士时空变化规律,发现此制度确实为北方人提供了考取进士的机会,但是总体来看,南方进士始终以压倒性优势胜于北方。主题河流图描绘了明代进士群体各省分期的分布特征,直观展现出永乐年江西进士数量占比最多,嘉靖年是各省进士数量最多的时期,进而可以更有效率地提出新命题供人文学者深思。在明代进士群体空间分布地图中,可以清晰知晓区域之间的分布差异,由粗粒度到细粒度呈现进士空间分布态势,发现不同省、府进士数量差异较大,且同一省、府内空间集聚效应明显,由此可以提出“进士空间分布是否受空间距离因素影响”这一思考,进一步可以量化距离要素,深入探讨距离对其影响程度,且“省力法则”也可以反映进士分布聚焦之因。

(2)在亲属关系研究中,从定量视角探求直系祖先对进士入仕率的影响,发现直系祖先无官职的占45%左右,据此可以推测寒门子弟考取进士入仕的机会也很大。另外本文仅以科举第一家族“常州庄氏”世系关系为例,用树状图绘制该进士在明代的亲属关系。虽然数据量有限,但为史学研究提供了技术视角下的研究范式,学者可以通过汇总相关史学数据、家谱等,从宏观或者历时性视角构建类似的亲属关系网络图,进而高效地梳理相关史料。

(3)在社会关系网络研究中,以“东林进士八君”为例,探讨该群体的网络结构属性、核心人物、小群体现象及空间分布特征,发现该群体的社会网络符合小世界效应及无标度分布特征,接近于现代社交媒体网

络结构,且“东林八君”社会关系人聚集之地即是进士荟萃之地,侧面反映了“东林八君”在当时社会的影响力之大。该研究为全面解读“东林八君”这一群体提供了一种思考方式,借此史学者可以利用明代进士数据从历时性视角构建类似网络结构,例如构建明代进士分期社会网络图,解析每个时期进士群体的社交网络关系,从中发现隐匿的规律。

参考文献:

[1] UNSWORTH J. What is humanities computing and what is not? [EB/OL]. [2020-05-03]. https://www.ideals.illinois.edu/bitstream/handle/2142/157/whatis.html.
[2] 徐力恒,陈静. 我们为什么需要数字人文[J]. 社会科学报,2017-08-24(5).
[3] 艾登. 可视化未来:数据透视下的人文大趋势[M]. 王彤彤,沈华伟,程学旗,译. 杭州:浙江人民出版社,2015.
[4] 中國歷代人物傳記資料庫(CBDB)[EB/OL]. [2020-06-03]. https://projects.iq.harvard.edu/chinesecbdb.
[5] BUSA F R. The annals of humanities computing: the index thomisticus[J]. Computers & the humanities,1980,14(2):83-90.
[6] MORETTI F. Distant reading[M]. London:Verso Books,2013.
[7] SCHICH M,SONG C,AHN Y Y,et al. A network framework of cultural history[J]. Science,2014,345(6196):558-562.
[8] JOCKERS M L. Macroanalysis: digital methods and literary history [M]. Urbana:University of Illinois Press,2013.
[9] DAKANG C. Identification of the authorship of the last 40 chapters of "a dream of red mansions" from the aspect of mathematical linguistic: discuss with Chen Bingzao[J]. Studies on "a dream of red mansions",1987(1):293-318.
[10] HORTON T,TAYLOR K,YU B,et al. Quite right, dear and interesting: seeking the sentimental in nineteenth century American fiction[J]. Digital humanities,2006:81-82.
[11] PLAISANT C,ROSE J,YU B,et al. Exploring erotics in Emily Dickinson's correspondence with text mining and visual interfaces [C]//Proceedings of the 6th ACM/IEEE-CS joint conference on digital libraries. New York:ACM,2006:141-150.
[12] LIU C L,HUANG C K,WANG H,et al. Mining local gazetteers of literary Chinese with CRF and pattern based methods for biographical information in Chinese history[C]//Proceedings of 2015 IEEE international conference on big data. Washington: IEEE, 2015: 1629-1638.
[13] 欧阳剑. 面向数字人文研究的大规模古籍文本可视化分析与挖掘[J]. 中国图书馆学报, 2016(2):66-80.
[14] FRANCO M. Network theory,plot analysis[C]//中国文艺理论学会.“文学与形式”国际学术研讨会暨中国文艺理论学会年会论文集. 南京:南京大学,2010:463-471.
[15] 赵薇. 社会网络分析与“《大波》三部曲”的人物功能[J]. 山东社会科学,2018 (9):11-25.
[16] 魏会洋,袁曦临. 社会网络分析在文学阅读研究中的适用性问题——以数字人文视角下的《白鹿原》人物关系阐释为例[J]. 新世纪图书馆,2019 (3):32-36.
[17] WETHERELL C. Historical social network analysis[J]. International review of social history,1998,43(S6):125-144.
[18] 严承希,王军. 数字人文视角:基于符号分析法的宋代政治网络可视化研究[J]. 中国图书馆学报,2018,44(5):87-103.
[19] NOVAK J,MICHEEL I,MELENHORST M,et al. HistoGraph——a visualization tool for collaborative analysis of networks from historical social multimedia collections[C]//2014 18th international conference on information visualisation. Paris: IEEE,2014:241-250.
[20] Visualizing networks of artistic ideas in history paintings in the seventeenth-century Netherlands[EB/OL]. [2020-06-03]. https://dev.clariah.nl/files/dh2019/boa/0136.html.
[21] KNOWLES A K. Emerging trends in historical GIS[J]. Historical geography,2005,33:7-13.
[22] GREGORY I N,HEALEY R G. Historical GIS: structuring, mapping and analysing geographies of the past[J]. Progress in human geography,2007,31(5):638-653.
[23] 丝绸之路历史地理信息开放平台[EB/OL]. [2020-06-03]. http://220.194.54.102/.
[24] 温永宁,闾国年,陈旻,等. 华夏家谱 GIS 的数据组织与系统架构[J]. 地球信息科学学报,2011,12(2):235-241.
[25] 王大学,陈熙,杨光辉. 基于 GIS 的中国古籍地理信息系统研究[J]. 复旦学报: 自然科学版,2016,55(6):684-688.
[26] 张萍. 地理信息系统(GIS)与中国历史研究[J]. 史学理论研究,2018(2):35-47.
[27] 胡迪,闾国年,江南,等. 地理与历史双重视角下的历史 GIS 数据模型[J]. 地球信息科学学报,2018,20(6):713-720.
[28] 黄鹏程.《列朝诗集小传》诗人地理分布的可视化呈现与阐释[J]. 图书馆论坛,2017, 37(5):47-54.
[29] 陈友乔. 明代广东进士的数量及时空分布[J]. 惠州学院学报,2019,39(5):43-51.
[30] 许娅. 明代贵州家族进士生平及著述丛考[J]. 教育文化论坛,2019,11(3):36-42.
[31] 沈登苗. 明清全国进士与人才的时空分布及其相互关系[J]. 中国文化研究,1999(4):59-66,3.
[32] 王红春. 明代进士家状研究[D]. 上海: 华东师范大学,2013.
[33] WASSERMAN S,FAUST K. Social network analysis: methods and applications[M]. Gambridge: Cambridge University Press,1994.
[34] 斯科特. 社会网络分析法[M]. 刘军 ,译. 重庆:重庆大学出版社,2007.
[35] 中华文明之时空基础架构 WMTS 服务[EB/OL]. [2020-06-03]. http://gis.sinica.edu.tw/ccts/.
[36] 邱进春. 明代江西进士考证[D]. 杭州:浙江大学,2006.
[37] 郭培贵,赵丽美. 明代广西进士人数及其地理分布考述[J]. 教育与考试,2010 (4):27-32.
[38] 何炳棣,徐泓. 明清社会史论[M]. 杭州:浙江大学出版社,2019.

[39] 弗里德曼,晓春. 中国东南的宗族组织[M]. 上海:上海人民出版社,2000.
[40] 郭培贵. 明代进士家族相关问题考论[J]. 求是学刊,2015(6):144－149.
[41] 叶可汗. 明代福建进士家族研究[D]. 长春:辽宁师范大学,2012.
[42] 中国历史大辞典(明史卷)[M]. 上海:上海辞书出版社,1995.
[43] WATTS D J, STROGATZ S H. Collective dynamics of 'small-world' networks[J]. Nature,1998,393(6684):440.
[44] BARABÁSI A L, ALBERT R. Emergence of scaling in random networks[J]. Science,1999,286(5439):509－512.
[45] NEWMAN M E J, GIRVAN M. Finding and evaluating community structure in networks[J]. Physical review E, 2004, 69(2): 026113.
[46] 王兆鹏,郑永晓,刘京臣. 借器之势,出道之新——"数字人文"浪潮下的古典文学研究三人谈[J]. 文艺研究,2019(9):79－88.

作者贡献说明：

邓君：提出研究思路，撰写论文，最终修订论文版本；
孙绍丹：论文撰写与修改；
王阮：论文数据采集；
宋先智：协助论文数据采集、论文校对；
宋雪雁：负责数据和论文校对。

Analysis of the Spatiotemporal Network Structure of the Imperial Examination Jinshi Groups of Ming Dynasty Under the Perspective of Digital Humanities

Deng Jun　Sun Shaodan　Wang Ruan　Song Xianzhi　Song Xueyan
School of Management, Jilin University, Changchun 130022

Abstract: [**Purpose/significance**] Digital technology is renewed and iteratively developed. Digital intelligent tools provide new research perspectives and paradigms for humanities research. [**Method/process**] By searching data of the elite group of Jinshi in Ming Dynasty in CBDB, using SNA and GIS, this paper drew the space-time network diagram of Jinshi group. [**Result/conclusion**] From 3 levels of the evolution law of time and space, the kinship network and the social relationship network, "distant reading" the historical phenomenon refracted behind the elite group of Jinshi in Ming Dynasty, helping humanists to more efficiently sort out historical facts and accurately grasp research propositions from complicated information. It highlights the huge potential of digital technology in the field of digital humanities.

Keywords: digital humanities　imperial Jinshi of Ming Dynasty　spatiotemporal network structure